

Unbiased test for a location parameter : case of logistic distribution

著者	Nogami Yoshiko
発行年	2000
著作の一部	Institute of Policy and Planning Sciences discussion paper series ~ no. 857
URL	http://hdl.handle.net/2241/621

INSTITUTE OF POLICY AND PLANNING SCIENCES

Discussion Paper Series

No. 857

Unbiased Test for a Location Parameter
-Case of Logistic Distribution-

by

Yoshiko Nogami

April 2000

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Unbiased Test for a Location Parameter (2).

---Case of Logistic Distribution---

By Yoshiko Nogami

Abstract.

In this paper we deal with the Logistic distribution with density

$$f(x|\theta) = \frac{e^{-(x-\theta)}}{\{1+e^{-(x-\theta)}\}^2}, \quad \text{for } -\infty < x < \infty$$

where $-\infty < \theta < \infty$. Based on a random sample X_1, \dots, X_n of size n from the density $f(x|\theta)$ we consider the problem of the testing the null hypothesis $H_0: \theta = \theta_0$ versus the alternative hypothesis $H_1: \theta \neq \theta_0$ for some constant θ_0 . We propose the test with the acceptance region derived from inverting the shortest confidence interval for θ_0 and check if this test is unbiased.

§1. Introduction.

In this paper we deal with Logistic distribution whose density is given as follows:

$$(1) \quad f(x|\theta) = \frac{e^{-(x-\theta)}}{\{1 + e^{-(x-\theta)}\}^2}, \quad \text{for } -\infty < x < \infty$$

provided that $-\infty < \theta < \infty$. Let X_1, \dots, X_n be a random sample of size n taken from the density $f(x|\theta)$. We find in Section 2 the confidence interval (C. I.) for θ with the shortest length using Lagrange's method. In Section 3 we consider the problem of testing the null hypothesis $H_0: \theta = \theta_0$ versus the alternative hypothesis $H_1: \theta \neq \theta_0$ for some constant θ_0 . We propose the test with acceptance region derived from inverting the shortest C. I. for θ_0 . Let α be a real number such that $0 < \alpha < 1$. When $n=2m+1$ with m a nonnegative integer, we show that our test is unbiased and of size α . But, when $n=2m$, because we use conventional device to get the C. I. for θ , we cannot show unbiasedness of our test. However, for large m our test becomes almost unbiased as the test in case of $n=2m+1$ shows.

Let $\hat{\theta}$ be the defining property.

§2. The Interval Estimation for θ .

Let X_1, \dots, X_n be a random sample of size n taken from the population with the density (1). We find the shortest C. I. for θ using Lagrange's method.

Let $n=2m+1$ with m a nonnegative integer, until (14). Let $X_{(i)}$ be the i -th smallest observation of X_1, \dots, X_n . We estimate θ by $Y = X_{(m+1)}$. To get the shortest C. I. for θ we first find the density of Y . Let $F(x|\theta)$ be the cumulative distribution function (c.d.f.) of X . Then, by (1) we get

$$(2) \quad F(x) = F(x|\theta) = \{1 + e^{-(x-\theta)}\}^{-1}, \quad \text{for } -\infty < x < \infty.$$

Hence, the density of Y is of form

$$(3) \quad g_Y(Y|\theta) = k(F(Y))^m(1-F(Y))^mf(Y|\theta), \quad \text{for } -\infty < Y < \infty.$$

where

$$(4) \quad k = \Gamma(2m+2) / \{\Gamma(m+1)\}^2.$$

Let α be a real number such that $0 < \alpha < 1$. Let r_1 and r_2 be real numbers such that $r_1 < r_2$. To find the shortest C. I. for θ at confidence coefficient $1-\alpha$ we want to minimize $r_2 - r_1$ under the condition that

$$(5) \quad P_\theta[r_1 < Y - \theta < r_2] = 1 - \alpha.$$

But, it follows by a variable transformation $W = F(Y)$ that

$$(6) \quad \text{the left hand side of (5)} = P_\theta[r_1 + \theta < Y < r_2 + \theta]$$

$$= P_\theta[F(r_1 + \theta) < W < F(r_2 + \theta)] = 1 - \alpha.$$

Hence, we want to minimize $r_2 - r_1$ under the condition (6). To do so we use Lagrange's method. Let λ be a real number and define

$$(7) \quad L = L(r_1, r_2; \lambda) = r_2 - r_1 - \lambda \left\{ \frac{F(r_2 + \theta)}{F(r_1 + \theta)} - \int_{F(r_1 + \theta)}^{F(r_2 + \theta)} h_W(w) dw - 1 + \alpha \right\}$$

where $h_W(w)$ is the density of W given by

$$(8) \quad h_W(w) = kw^m(1-w)^m, \quad \text{for } 0 < w < 1$$

where k is given by (4). The right hand side of (8) is the probability density function(p.d.f.) of Beta distribution $\text{Beta}(m+1, m+1)$ with $(m+1, m+1)$ degrees of freedom. Then, by Lagrange's method we have that

$$(9) \quad \begin{cases} \partial L / \partial r_1 = -1 + \lambda h_w(F(r_1 + \theta)) f(r_1 + \theta | \theta) = 0 \\ \partial L / \partial r_2 = 1 - \lambda h_w(F(r_2 + \theta)) f(r_2 + \theta | \theta) = 0 \end{cases}$$

By (9) we get that

$$(10) \quad h_w(F(r_1 + \theta)) f(r_1 + \theta | \theta) = h_w(F(r_2 + \theta)) f(r_2 + \theta | \theta) \quad (= \lambda^{-1}), \quad \forall \theta.$$

Taking

$$(11) \quad F(r_1 + \theta) = \beta(a/2) \quad \text{and} \quad F(r_2 + \theta) = 1 - \beta(a/2)$$

where $\beta(a/2)$ is given by

$$(12) \quad \int_0^{\beta(a/2)} h_w(w) \, dw = a/2,$$

we obtain by (2) that $r_1 = -r_2 = -r$ where

$$(13) \quad r = F^{-1}(1 - \beta(a/2)) - \theta = \ln[\{1 - \beta(a/2)\} / \beta(a/2)].$$

We also have $h_w(F(-r + \theta)) = h_w(F(r + \theta))$ and $f(-r + \theta | \theta) = f(r + \theta | \theta)$ with r given by (13).

Thus, (10) and (6) are satisfied for $r_1 = -r_2 = -r$ with r given by (13). Therefore,

the shortest C. I. for θ at confidence coefficient $1 - \alpha$ is given by

$$(14) \quad (Y - r, Y + r) = (Y - \ln[\{1 - \beta(a/2)\} / \beta(a/2)], Y + \ln[\{1 - \beta(a/2)\} / \beta(a/2)]).$$

Let $n = 2m$. This time we estimate θ by $Y = X_{(m)}$. In the similar way to the above we get the density of Y

$$(15) \quad g_Y(Y | \theta) = k_1 (F(Y))^{m-1} (1 - F(Y))^m f(Y | \theta), \quad \text{for } -\infty < Y < \infty$$

where

$$(16) \quad k_1 = \Gamma(2m+1) / \{\Gamma(m)\Gamma(m+1)\}.$$

Putting $W=F(Y)$ we minimize $r_2 - r_1$ under the condition (6). However, since the density of W is now of form

$$(17) \quad h_1(w) = k_1 w^{m-1} (1-w)^m, \quad \text{for } 0 < w < 1$$

which is the p.d.f. of the Beta($m, m+1$) distribution with k_1 defined by (16), it is difficult to get exact values for $F(r_i + \theta)$, $i=1, 2$ which satisfy

$$(18) \quad h_1(F(r_1 + \theta))f(r_1 + \theta | \theta) = h_1(F(r_2 + \theta))f(r_2 + \theta | \theta).$$

Hence, we use conventional values for $F(r_i + \theta)$, $i=1, 2$. Those are

$$(19) \quad F(r_1 + \theta) = \beta_{m, m+1}(\alpha/2) \quad \text{and} \quad F(r_2 + \theta) = 1 - \beta_{m+1, m}(\alpha/2)$$

where $\beta_{m, m+1}(\alpha/2)$ and $\beta_{m+1, m}(\alpha/2)$ are respectively determined by

$$(20) \quad \int_0^{\beta_{m, m+1}(\alpha/2)} h_1(w) dw = \alpha/2 = \int_0^{\beta_{m+1, m}(\alpha/2)} k_1 w^m (1-w)^{m-1} dw.$$

Thus, r_1 and r_2 are respectively given by

$$(21) \quad \begin{cases} r_1 = F^{-1}(\beta_{m, m+1}(\alpha/2)) - \theta = -\ln[\{1 - \beta_{m, m+1}(\alpha/2)\} / \beta_{m, m+1}(\alpha/2)] \\ r_2 = F^{-1}(\beta_{m+1, m}(\alpha/2)) - \theta = \ln[\{1 - \beta_{m+1, m}(\alpha/2)\} / \beta_{m+1, m}(\alpha/2)] \end{cases}$$

Therefore, the C. I. for θ at confidence coefficient $1-\alpha$ is

$$(22) \quad (Y - r_2, Y - r_1),$$

where r_1 and r_2 are determined by (21).

In the next section we check if the tests with the acceptance regions derived from inverting the C. I.'s (14) for $n=2m+1$ and (22) for $n=2m$, respectively are unbiased and of size α .

§3. Two-Sided Test for θ .

In this section we consider the problem of testing the null hypothesis $H_0: \theta = \theta_0$ versus the alternative hypothesis $H_1: \theta \neq \theta_0$ for some constant θ_0 . We propose the two-sided test with the acceptance region derived from inverting the shortest C. I. for θ_0 . When $n=2m+1$ we show that our test is unbiased and of size α . When $n=2m$ our test is not unbiased because of usage of conventional method for constructing the C. I. for θ .

Let $n=2m+1$. As in Section 2 we define $Y = X_{(m+1)}$. By inverting the shortest C. I. (14) for θ_0 our test is to reject $Y \in (-\infty, \theta_0 - r] \cup [\theta_0 + r, +\infty)$ and to accept H_0 if $Y \in (\theta_0 - r, \theta_0 + r)$ where r is given by (13). Now, we show that this test is unbiased and of size α .

Let y_1^0 and y_2^0 be real numbers depending on θ_0 such that $y_1^0 < y_2^0$. Define $\psi(\theta)$ by

$$\begin{aligned} (23) \quad \psi(\theta) &= P_\theta [Y < y_1^0 \text{ or } y_2^0 < Y] \\ &= 1 - \int_{y_1^0}^{y_2^0} g_Y(Y|\theta) \, dY \end{aligned}$$

where $g_Y(Y|\theta)$ is defined by (3). To get unbiased size- α test with the acceptance region (y_1^0, y_2^0) we choose y_1^0 and y_2^0 which satisfy

$$(24) \quad \psi(\theta_0) = 1 - P_{\theta_0} [y_1^0 < Y < y_2^0] = \alpha$$

and minimize $\psi(\theta)$ at $\theta = \theta_0$; namely

$$(25) \quad \left. \frac{d\psi(\theta)}{d\theta} \right|_{\theta=\theta_0} = g_Y(y_2^0|\theta_0) - g_Y(y_1^0|\theta_0) = 0.$$

We consider the test with the acceptance region $(\theta_0 - r, \theta_0 + r)$. Since from the construction the equality (10) with $r_1 = -r$, $r_2 = r$ and $\theta = \theta_0$ is satisfied, we obtain by (3) and (8) that $g_Y(\theta_0 - r | \theta_0) = g_Y(\theta_0 + r | \theta_0)$; (25) is satisfied for y_1^0 and y_2^0 replaced by $\theta_0 - r$ and $\theta_0 + r$, respectively. (24) with y_1^0 and y_2^0 replaced by $\theta_0 - r$ and $\theta_0 + r$, respectively is the same as (5) except for θ , r_1 and r_2 replaced by θ_0 , $-r$ and r , respectively. Therefore, our test with the acceptance region $(\theta_0 - r, \theta_0 + r)$ is unbiased and of size α .

Let $n=2m$. As in Section 2 we define $Y \stackrel{\Delta}{=} X_{(m)}$. Again, by inverting the C. I. (22) for θ_0 our test is to reject H_0 if $Y \in (-\infty, \theta_0 + r_1] \cup [\theta_0 + r_2, +\infty)$ and to accept H_0 if $Y \in (\theta_0 + r_1, \theta_0 + r_2)$ where r_1 and r_2 are given by (21). In this case our test depends on the conventional values for $F(r_i + \theta)$, $i=1, 2$. Hence, we have that $g_Y(\theta_0 + r_1 | \theta_0) \neq g_Y(\theta_0 + r_2 | \theta_0)$. Furthermore, (24) with y_1^0 and y_2^0 replaced by $\theta_0 + r_1$ and $\theta_0 + r_2$, respectively is the same as (5) except for θ replaced by θ_0 . Therefore, our test is still of size α , but not unbiased. However, for large m our test becomes almost unbiased as the test in case of $n=2m+1$ shows.